DOCUMENT RESUME

ED 126 717 FL 007 820

AUTHOR Gazdar, Gerald
TITLE Quantifying Context. York Papers in Linguistics, No. 6.
INSTITUTION York Univ. (England). Dept. of Language.
PUB DATE Mar 76
NOTE 14p.

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS Data Analysis; Descriptive Linguistics; Generative Grammar; Graphs; Humor; *Linguistic Theory; Morphology (Languages); *Probability Theory; *Scientific Methodology; Semantics; Sociolinguistics; *Statistical Analysis; Statistical Data; *Structural Analysis; Transformation Theory (Language)
IDENTIFIERS *Probabilistic Grammar

ABSTRACT

This is a tongue-in-cheek rendering of an imaginary linguist's attempt to prove that a grammatical rule is not in fact optional, but conditioned. Through manipulation of the data, use of computers, etc., he succeeds in "proving" this. He also "proves" that this conditioning is contextual; that the rule which exhibits it is essentially probabilistic; and that hyperbolic tan functions are a feature of the competence underlying the human language faculty. An alternative, much simpler hypothesis is presented, but rejected on the grounds that, although it is compatible with the data, it is not "empirical" enough. Assumptions implicit in variable rule descriptions (VRD's) are examined, with the conclusion that the VRD approach has unfortunate methodological consequences for linguistic theory. It is further concluded that probabilistic grammar is not a theory. (Author/AM)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ED126717

FL007820

QUANTIFYING CONTEXT*

Gerald Gazdar

(School of Social Sciences,
University of Sussex)

This paper consists of a story with a moral. In order to preserve the continuity of the story I shall put quite a bit of the moral into the footnotes. Those with no taste for morals can, of course, skip the footnotes.

This story is about some linguistic data. And, because the point of the story is not to clue you in to one or two facts about an obscure language, I'm going to take some liberties with the facts. What matters is that the story *could* be true, not whether it is or not. It is, perhaps,[1] the first ever sociolinguistic shaggy dog story.

There's a language, which I'll call Tzotzil[2] for which someone once wrote a Mickey Mouse generative grammar.[3] Among the rules of this grammar is something which we can represent thus:[4]

(1) WHERE → $\begin{Bmatrix} k'al \\ teno \end{Bmatrix}$

In other words the underlying where-question morpheme has two different realizations in free variation. And you'll have to take my word for the fact that the rule isn't syntactically or phonologically conditioned.

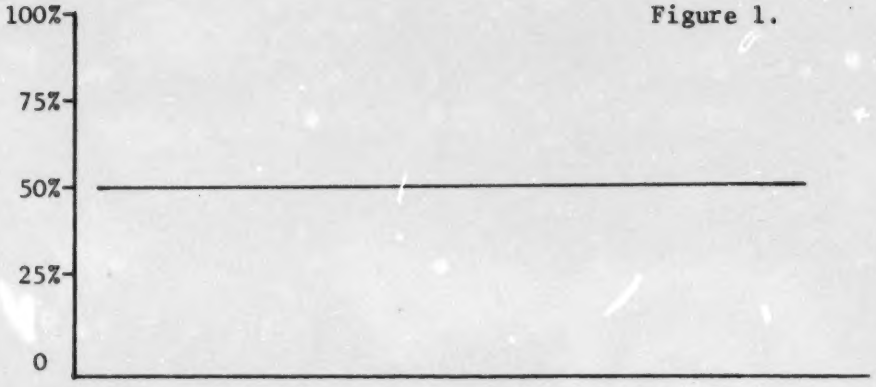
As is well known, one thing hard-nosed sociolinguists don't wear is braces. So, when one of their number, whose name I won't mention,[5] was browsing through this grammar in his bath one morning, he winced a bit when he saw (1). Clearly there was only one thing he could do -- pack up four research assistants and five Uhers and parachute into Tzotzil-land to check the rule out.

Out in Tzotzil-land he and his baggage laboriously record every occurrence of *k'al* and *teno* and for each occurrence they note down the age of the speaker, where he was standing, his pupil dilation, the time of day, the home address of him and his addressee, the air temperature, the height of the speaker's spouse, and one or two other things which I'd rather not mention.

Meanwhile, back in the U.S.A., the computer is busy thinking about the data that our linguist is diligently radioing back to it. As a result of its cogitations it plots a number of graphs, some of which will be reproduced below. In the first one [Figure 1. overleaf], the y-coordinate ranges from 1-100 and represents the percentage of *teno* occurrences relative to total *k'al* and *teno* occurrences. Obviously 100 less whatever the value of y is will give you the percentage of *k'al* occurrences.

(2) $y = \frac{100n}{n+m}$ 2 where n = number of *teno* occurrences
where m = number of *k'al* occurrences

Figure 1.

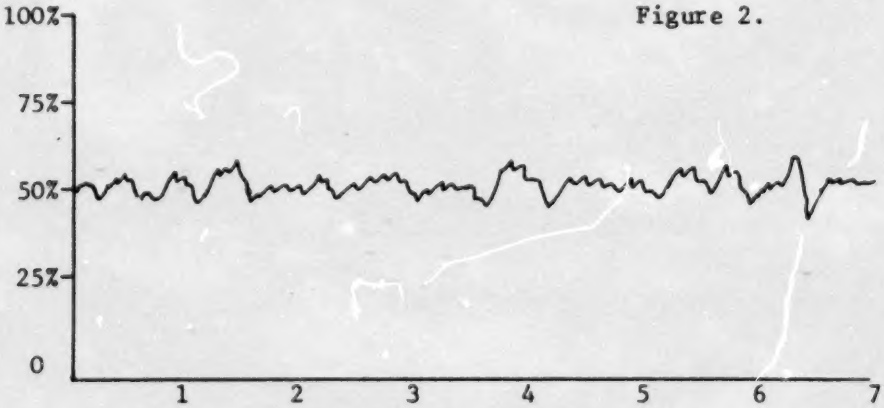


Pupil dilation

Along the abscissa we have pupil dilation. Now, as the reader well knows, indeed as anyone knows who has read Labov (1970), pupil dilation is an objective empirical measure of the degree of formality of a conversational situation. You may be wondering how Labov ever discovered that,[6] but try and put such unworthy thoughts out of your mind. However, as you can see, this graph actually isn't very exciting. It seems that pupil dilation, a.k.a. 'degree of formality', is completely irrelevant to relative *k'al* and *teno* usage.

In the next graph the y-coordinate stays the same but along the abscissa we have the home address and occupation of the speaker ranged along a seven point scale. Some of my more naive readers may be a little puzzled as to how we match such obviously multidimensional data to a single linear seven point scale. Let me fill you in on the sophisticated technique involved:[7] what you do is award points. So if the subject lives in the Tzotzil-land equivalent of Park Lane they get 7 points for home address, whereas if they live in the Tzotzil-land equivalent of Clapham they only get 2 points. Likewise if the subject is the Tzotzil equivalent of Vice-Chancellor then they get 6 points whereas if they are the Tzotzil equivalent of a cleaner they only get 1 point. For any given subject you just tot up their scores, divide by the number of dimensions you first thought of, and plot them on a graph. I'm sorry if this digression has bored those who knew how to do it all along.

Figure 2.

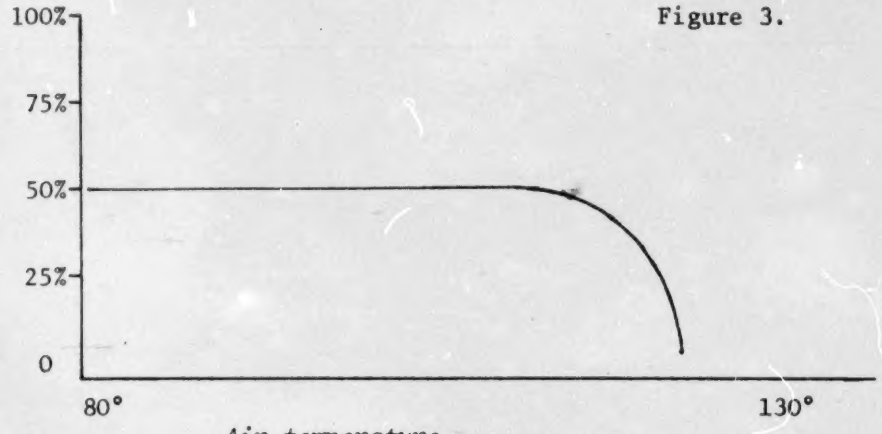


Occupation and home address

As you can see, this graph is a little more interesting than the last one but our linguist was terribly disappointed with it. I should have mentioned that the computer not only prints out these pretty graphs but it also associates with each a string of figures. And in this case the figures tell us that the curve is just a random distribution around the 50% line, a distribution which is of no significance to man or beast. Our linguist is distressed at this because his previous experience of this variable had led him to believe that almost anything which linguists had called optional was, in reality, strongly correlated with it.

Disappointed, but undeterred, our linguist continued to examine the graphs provided by the computer, despite the fact that the two variables which were, in his view, the most promising, had let him down. Being a good empiricist he was naturally willing to examine the correlation between any pair of variables whatsoever.[8] The next graph to emerge from the line-printer turned out to be very puzzling indeed:

Figure 3.



In this graph the abscissa represents air temperature. As you can imagine this curve had our linguist seriously worried. He was after all, a *socio*linguist not a physiological linguist. Whatever else he wanted he didn't want the honour of being the first linguist to discover an 'optional' rule that was conditioned by air temperature.[9] Accordingly he took a close look at the string of figures associated with this graph and discovered that although the curve was significant it wasn't really very significant. There was a one-in-ten possibility of its having happened by chance. Our linguist put this graph on one side, consoling himself with the thought that he was probably the victim of one of those statistical freaks that even the most rigorous of scientists must sometimes fall prey to.

Several more graphs came out but they all followed the pattern of the first two we looked at so I won't bore you with the details. The linguist was beginning to get rather anxious -- maybe the rule really was optional. The final graph to emerge swept all these worries away -- here at last was something exciting; not only did it have a pretty shape but also the figures associated with it made it virtually certain that it had not arisen through chance.

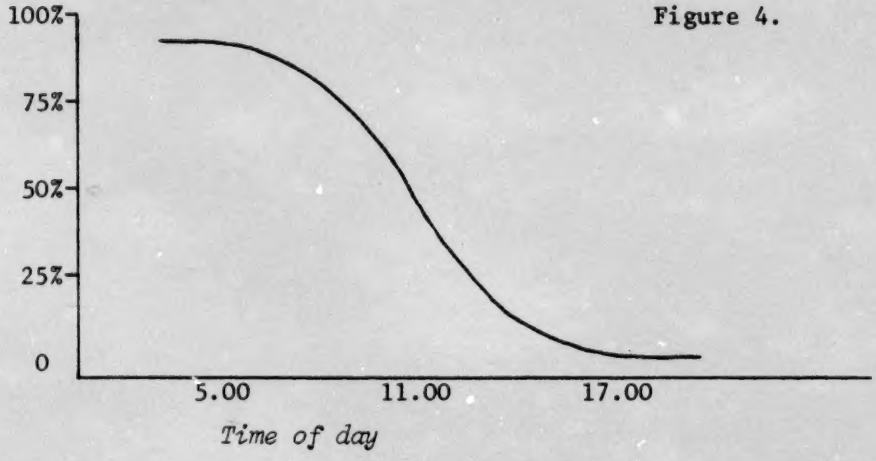


Figure 4.

The abscissa of this graph represents the time of day -- the curve turns at about 11.00 a.m. At last we have the proof that the *k'al / teno* alternation is not optional. Our linguist looked at this graph and the first hypothesis that sprang to his mind went as follows: there is an abstract competence-type rule which says use *teno* before 11.00 and use *k'al* after 11.00. Furthermore the output of this competence rule is obscured by performance factors such as memory limitations, slips of the tongue, distractions, false starts, shifts of attention and interest, and faulty watches sold to the Tzotzil-landers by unscrupulous Japanese salesmen.[10] In the absence of these performance factors the competence rule would appear thus:[11]

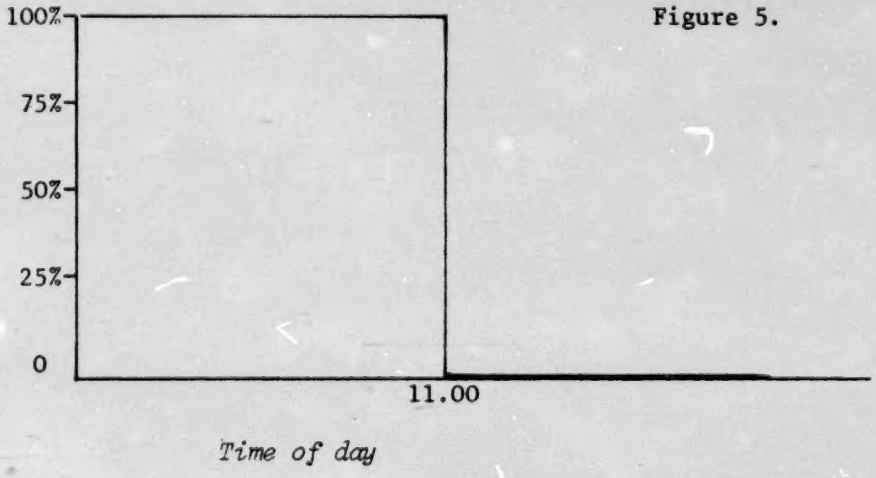


Figure 5.

On further reflection our linguist decided that this hypothesis wouldn't do. There were three reasons for this, the last being the most compelling, although only the first two get mentioned in his forthcoming paper.

The first is that the curve isn't quite what you'd expect on the most plausible model of performance discrepancy. If speakers switched from *teno* to *k'al* as soon as they thought it was 11.00, but sometimes got the time wrong then the curve would look like Figure 6 and not like Figure 4:

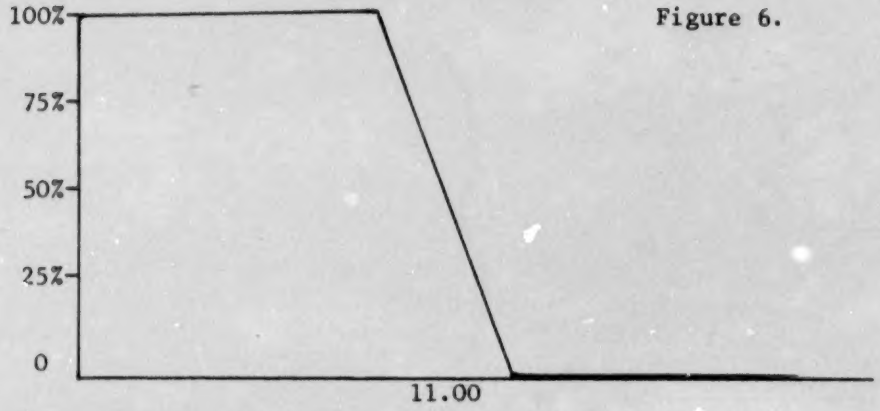


Figure 6.

The second reason is that if you look at the individual data, rather than at the grouped data, then you discover that the curve shown in Figure 4 is replicated for individual speakers and isn't just a product of their summation.

The third reason -- the unmentionable one -- lay in something one of his research assistants confided to him as they were on their way to a Tzotzil Bar Mitsvah one day, soon after they had formulated the 11.00 competence hypothesis. This research assistant had apparently overheard one of their subjects use the *k'al* form at 6.00 one morning. The research assistant approached the subject and asked him to repeat what he had just said. The subject repeated his utterance and again used the *k'al* form. The research assistant then asked the subject whether he realized that he had made a mistake and whether he knew what the time was. The subject replied that he wasn't aware of having made any mistake and that he thought the time was about 6.00 a.m. Needless to say, the research assistant came away a puzzled and worried man.

Of course an anecdote like that has no place in an empirical theory of language and our linguist rightly disregards it completely when he justifies his rejection of the 11.00 competence hypothesis.[12]

The linguist's new hypothesis took the form of a variable rule, a rule which, like other rules of generative grammar is to be interpreted as a part of individual competence.[13] This rule looks like this: [14]

(3) WHERE $\longrightarrow \begin{bmatrix} \langle teno \rangle_{n}/t \\ \langle k'al \rangle_{1-n}/t \end{bmatrix}$ where n = f(t) and t is the time of the utterance.

Now that may look very fancy, but it doesn't help us too much unless we know what f is. Luckily that question is easy to answer -- I expect most readers who have given Figure 4 even a cursory glance will already have guessed it. Yes, of course, Figure 4 is none other than a close approximation to our old friend the common or garden negative hyperbolic tan curve. So we have that:[15]

(4) $f = \lambda x(a - \tanh(bx - c))$
where a, b and c are constants.

Our linguist got the computer to fiddle around with the constants a bit and came up with best-fit values for a, b and c. Then he tested the match between the curve generated by his variable rule and that plotted as a result of his observations. Amazingly the match is very close and the possibility of its occurring by chance quite negligible.

Before bringing this saga of research virtuosity to a close let me just summarize the achievements of this linguist. He has shown that yet another rule, previously thought optional, is actually conditioned. He has shown that this conditioning is contextual and that the rule which exhibits it is *essentially* probabilistic and not just an absolute rule obscured by performance factors. And above all, innatists rejoice, he has shown that hyperbolic tan functions are a feature of the competence underlying the human language faculty.

I might mention, in passing, that an additional graph provided by the computer also enabled our linguist to explain the very puzzling facts he had uncovered concerning the relation between where-questions and air-temperature. The computer had provided him with a plot of air-temperature against time which looked like this:

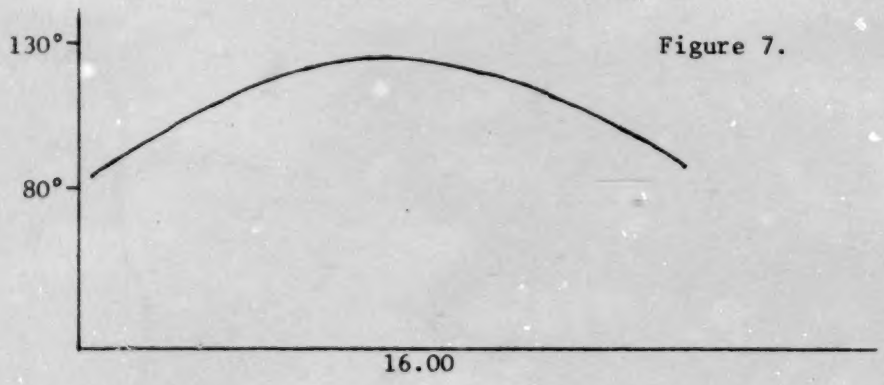


Figure 7.

This shows that the hottest part of the Tzotzil day is about 4.00 in the afternoon. If we compound this curve with our time-dependent variable rule then the graph (Figure 3) that we had before, appears quite straightforwardly. Here, as if he needed it, was independent confirmation of our linguist's hypothesis.

Now I'm sure that the reader has found our linguist's second hypothesis utterly compelling as an explanation of the curious facts about *teno* and *k'al*. However I feel that it is only fair to point out that there is an alternative hypothesis available, if only to illustrate the immense superiority of the theory just described. The trouble with this alternative hypothesis is that it isn't very empirical. In fact it isn't at all empirical. And it invokes non-behavioral concepts like 'meaning'.[16] Worse still it doesn't make any very precise[17] predictions about the data.

What predictions it does make seem more or less compatible with the data as I've described it. Anyway let me just sketch this hypothesis for you so you can see how vague, silly and methodologically unsatisfactory it is.

The silly hypothesis runs as follows: the words *k'al* and *teno* mean slightly different things.[18] Both are basically equivalent to the English word 'where' but *teno* carries in addition the presupposition that the addressee hasn't been very far away whilst *k'al* carries the presupposition that the addressee has been some distance away. And that's all we need to say *linguistically*. However there are one or two nonlinguistic facts which impinge on things. Before 11.00 people in Tzotzil-land haven't had a chance to go any great distance away and return so where-questions addressed to them mostly take the *teno* form. After 11.00 people start filtering back from the fields and the *k'al* form takes over in questions addressed to these persons. And that's it. That's the hypothesis. Of course it's silly to think that one could still get away with explanations like that now that linguistic theory is at liberty to posit variable rules and internalized hyperbolic tan functions.

* * * *

Let us examine the assumptions implicit in variable rule descriptions (VRDs hereafter) in rather more detail.[19] VRDs involve the linguist in a complete change of data base: they purport to offer an explanation, not for the form of sentence-types (as a conventional grammar does) but instead for the form of sentence-tokens, only the latter count as data for a VRD. Furthermore a VRD cannot employ as data the form of a particular sentence token (as a conventional grammar can) but can only refer to the set of forms found in a set of sentence-tokens.[20]

The explanandum of, and data for, conventional grammars is, in large measure, a set of judgements about well-formedness. These judgements are irrelevant to the probabilistic components of a VRD because the latter can only make statements about the relative frequency of occurrent forms. It might be argued that VRDs are about the well-formedness of texts but this view cannot be maintained. Consider the synonymous expressions '---' and '-' in the propositional calculus:[21] the former is much less common in proofs that the latter. Suppose we examine a large number of proofs and find that the ratio of former to latter is 1:3 and so we add a variable rule to the syntax of the propositional calculus to reproduce this distribution. What happens if we are then shown a text consisting of a set of proofs in which the occurrence ratio of '---' to '-' is 2:1? Are we to say that this text is ill-formed? Clearly not, all we can say is that it has a nonstandard occurrence ratio of '---' and '-'. VRDs conflate two distinct types of fact,[22] facts about well-formedness and facts about occurrence, 'all within the same notational and theoretical framework' (Cedergren and Sankoff 1974:334).

This conflation has a number of unfortunate methodological consequences for linguistic theory. In the first place VRDs make single counterexamples irrelevant. How could one produce a counterexample to the VRD proposed above for *k'al* and *teno* usage? The alternative semantic hypothesis is, on the other hand, clearly falsifiable by a single counterexample.[23] Worse still, imagine a linguist, sympathetic to VRDs, who is

confronted with a set of counterexamples to one of his *non*-probabilistic hypotheses: he has only to convert the hypothesis into an appropriately framed VRD in order to actually *predict* the 'counterexamples'. For example, suppose Schegloff (1968) had anticipated Labov's (1970) suggestion to employ variable rules in the analysis of dialogue data. Then, presented as he was with one counterexample to a rule[24] that held for the other 499 cases in his corpus, Schegloff could simply have weighted his rule to apply in 99.8% of cases. As it was Schegloff took the counterexample seriously[25] and postulated a more general non-probabilistic analysis which subsumed the entire corpus.

Perhaps the ultimate *reductio* of the VRD approach will come about when it confronts the problem of the relative distribution of synonymous expressions. Where these expressions are both listed in the lexicon we may simply add a probability to the entries to give us such pairs as ⟨OCULIST: [.997]⟩, ⟨EYE-DOCTOR: [.003]⟩. However, if we are going to do that, and the logic of the VRD approach makes it mandatory that we do, then we are faced with an insuperable problem when we are called upon to account for the relative distribution of transformationally unrelated but synonymous expressions, one or more of which is not to be found in the lexicon. Thus it would be impossible to formulate a VRD to account for the relative distribution of *perfect tense* and *passé historique* in French[26] or for the relative distribution of the expressions 'sister-in-law' and 'spouse's sister or brother's wife'.

The type of linguistics caricatured in this paper has, in effect, rediscovered discovery procedures. Just as distributional data analysis led to phrase-structure grammar and the taxonomic phoneme, so statistical data analysis leads to probabilistic grammar. Doubtless horoscopic data analysis leads to astrological grammar. Chomsky's discussions[27] of the methodological fallacy inherent in the discovery procedure procedure for theory construction are too well known to require repetition here.

There are two possible responses to these remarks. One is to say that VRDs, despite the grandiose claims made for them,[28] should not be assessed as theoretical statements at all but rather as descriptions meeting the criterion of observational adequacy, that is as descriptions which present the observed primary data correctly and thus achieve the lowest level of success for grammatical description. But, as Chomsky has pointed out, 'what data is relevant is determined in part by the possibility for a systematic theory, and one might therefore hold that the lowest level of success is no easier to achieve than the others.... The problem of deciding what data is valuable and to the point is not an easy one. *What is observed is often neither relevant nor significant*, and what is relevant and significant is often very difficult to observe, in linguistics no less than in the freshman physics laboratory, or, for that matter, anywhere in science.' (Chomsky 1964:28, fn.1 [my emphasis]).

That quantifying context should, in present circumstances, be a pointless and misleading exercise is quite unsurprising in the light of standard scientific methodology. For, as Chomsky implies in the quotation above, a component of that methodology is that theories determine their own data. This principle is no news to most linguists:[29] the linguist with a syntactic theory will study strings which are very different from

those investigated by the linguist with a phonetic theory. Equally, no linguist is interested in just any set of unacceptable strings but only those that confirm, disconfirm or decide between linguistic theories.

The second response is to claim that probabilistic grammar is indeed a theory and not just a reification of a particular type of discovery procedure. And to claim in addition that it is a theory which specifies that quantifiable aspects of context count as potentially relevant data.[30] But probabilistic grammar, as formulated by VRD proponents, is not a theory. Fancy statistics which summarize the distribution of items in a sample do not constitute an explanation of that distribution.[31] To think that they could do so, is to trade on a misconception of the role of statistics in science: the accepted role of statistics, even in psychology, is to test an explanation, not to *be* the explanation.[32] The first word in the phrase 'significant at the .01 level' does not mean the same as the second word in the phrase 'linguistically significant generalization' -- to proceed as if it did is to be guilty of gross equivocation. Significance in the first sense is no guarantee of significance in the second.

NOTES

* I am grateful to Mike Oakes for some discussion of statistical matters.

1. A rather cynical syntactician at the conference, on hearing me make this claim in conversation, responded with 'Aren't they all'.

2. There is actually a language called Tzotzil which does have where-question morphemes *k'at* and *teno*. It is interestingly discussed in Sarles 1970 to which I owe the inspiration for the present discussion. The latter should not however be construed either as a contribution to the study of the real language Tzotzil, or as a caricature of my source. Reference to Sarles 1970 will rapidly serve to dispel both these construals.

3. To the best of my knowledge this statement is false.

4. Of course there is an alternative formulation in which we assume one of the two forms to be underlying and have an optional rule which converts it into the other form. Nothing said below is in the least affected by my choice of formulation.

5. After this paper was given a leading British sociolinguist criticized it on the grounds that it should have named names and taken a specific piece of work as its target. The reason for not having done so is straightforward: the criticisms advanced are of a quite general methodological nature, to have directed them at a specific piece of work would have obscured the fact that they apply to a whole paradigm of research.

6. Cf. Chomsky 1964:56f.

7. For more serious discussions of the ontological absurdity of this kind of approach see Cicourel 1964, 1972; Gazdar 1976a and Wilson 1970.

8. Incredible as it may seem, the definition of empiricism implicit in this remark is also implicit in the work done by many persons referred to as 'social scientist'. It receives classic expression in the statistical technique known as *factor analysis*. 'There is surely no reason why the linguist must necessarily limit himself to "the study of phenomena and their correlations", avoiding any attempt to gain insight into such data by means of an explanatory theory of language.' (Chomsky 1964:99).

9. I do not mean to imply that rules of this type could not exist. A phonetic rule which, say, varied the height of back vowels according to atmospheric pressure does not sound, *a priori*, very unreasonable. What makes this type of rule so ludicrous in the present context is the fact that we are dealing with the distribution of lexical items.

10. This list is mostly derived from the lists given in Chomsky 1965.

11. It is not my intention to suggest that such a time-sensitive linguistic rule is intrinsically absurd. In the discussion following the paper C.A.Ferguson drew attention to the rule for the use of the Moroccan-Arabic word for 'needle' which appears to be of exactly this form for many speakers.

12. For more serious remarks on the inappropriacy of behaviorist restrictions on mentionable linguistic data see Gazdar 1976b. A behaviorist methodology is, of course, not a necessary concomitant of VRDs (Variable Rule Descriptions), or vice versa. However, being equally symptomatic of a misplaced empiricism, their cooccurrence in this instance should be unsurprising. Cf., e.g. the behavioral approach adopted by Labov (1973) in his attempt to come up with a probabilistic account of the meaning of the word 'cup'. At least some sociolinguists do not feel bound by a behaviorist methodology: Gillian Sankoff (personal communication) apparently only posits a VRD when she has failed to elicit a meaning difference from informant intuitions. This order of procedure seems to carry an implicit admission of the inferiority of VRDs relative to more traditional forms of linguistic explanation.

13. The latter half of this sentence is lifted, almost word for word, from Cedergren and Sankoff (1974:335).

14. The notational conventions employed are discussed in Cedergren and Sankoff 1974. The contextual variables most often found in such rules are 'formality' and 'class'. I have deliberately chosen to use *time* as the variable because its ontological status as something oriented to by speakers is that much more secure than that of the more commonly found variables (cf. footnotes 7 and 11 above). The point is that VRDs can be criticized even when they employ relatively sensible variables.

15. Readers unfamiliar with the lambda operator are directed to Feys and Fitch 1969.

16. I'm assuming that Quine's 'innocent and uncontroversial notion of stimulus meaning' (Quine 1972) is inadequate to the present task.

17. Clearly 'precise' means 'numerically precise'.

18. 'In general, it seems reasonable to regard an item as meaning-bearing just in case selection of it is subject to an optional rule... Where the grammar provides for an optional choice, it makes sense to search for the conditions under which it is appropriate to make this choice (this being one aspect of the study of meaning).' (Chomsky 1964:36-7, fn.7)

19. The discussion below is complementary to that found in Bickerton 1971, 1973a, 1973b and Gazdar 1976a.

20. Cf. Sanders 1974:7, fn.3.

21. Although '---' and '-' can correctly be said to be synonymous in the propositional calculus, they typically have a very different function in proofs. Facts like that can only be obfuscated by a VRD.

22. For a most useful discussion of this distinction see Dretske (1974: 3ff.). The fact, if it is one, that the string *cat on sat mat* never occurs *may* be evidence for the fact that the string is ill-formed. The grammar will explain its ill-formedness but its non-occurrence will only be explained by the grammar *taken together with* some rather strong psychological assumptions. The fact of its non-occurrence and the fact of its ill-formedness are thus facts of a quite different type.

23. E.g. A sees B leave the room and come back a couple of minutes later. A then asks B where B has been and uses the *k'al* form.

24. 'A first rule of telephone conversation, which might be called a "distribution rule for first utterances" is: *the answerer speaks first*.' (Schegloff 1968)

25. 'While indeed there is only one such violation in my data, its loneliness in the corpus is not sufficient warrant for not treating it seriously. Two alternatives are open. We might focus exclusively on this case and seek to develop an analysis particular to it that would account for its deviant sequencing. This would constitute an ad hoc attempt to save the distribution rule, using a technique commonly used in sociology -- deviant case analysis. Alternately, we might reexamine the entire corpus of materials seeking to deepen our understanding of the opening sequencing. We might ask: Is this best treated as a deviant case, or would a deeper and more general formulation of the opening sequencing reveal properties of the initiation of talk that the distribution rule glosses over. Analysis of the case reveals that the distribution rule, while it holds in most cases, is in fact best understood as a derivative of more general rules.' (Schegloff 1968)

26. I owe this example and the argument it illustrates to Dick Hudson (personal communication).

27. E.g in Chomsky 1962.

28. 'The power of this approach lies in the uniquely well-defined and economical relationship which it posits between competence and linguistic performance, analogous to that between a probability distribution and a sample, or between a model and a simulation. This relationship not only integrates generative and behavioral aspects in an elegant way, but is also easily operationalized so as to provide consistent and reproducible results. The utility of a theory containing variable rules is magnified many times as a consequence of the ease and naturalness with which it extends from purely linguistic applicability to the domains of sociolinguistics, stylistics, dialectology, and language change, individually or in combination.' (Cedergren and Sankoff 1974:353)

29. It is one of the themes of a recent symposium (Cohen 1974) on explaining linguistic phenomena; see especially the papers by Bever (1974), Dougherty (1974), Dreŕske (1974) and Whitaker(1974). It can also be found throughout Chomsky's methodological writings, e.g. 'The prevailing attitude in the sciences is to regard data as of interest primarily insofar as it has bearing on the choice among alternative theories, and to search for data, however exotic, that will be crucial in this sense.' (Chomsky 1964:98-9)

30. 'Once accepted and incorporated into description, variability can be made a function not only of the presence or absence of linguistic elements but also can be constrained by extralinguistic factors, *all within the same notational and theoretical framework.*'(Cedergren and Sankoff 1974:334 [my emphasis])

31. By way of analogy consider a complex algebraic equation which, when plotted on a graph, gave us an outline map of the U.K. There is no sense in which such an equation could be said to *explain* the shape of the U.K. although it might, e.g., be a useful way of storing the shape in a computer. [This analogy is not original to this paper but I can no longer locate a source for it, so none is given.]

32. The only counterexample I know of to this claim in the 'hard' sciences lies in the role of statistics in Quantum Mechanics. Persons wishing to draw parallels between the latter and linguistics are welcome to try and do so.

REFERENCES

Bailey, C.-J. & Shuy, R. (eds 1973) *New Ways of Analyzing Variation in English.* Washington D.C., Georgetown U.P.

Bever, T.G. (1974) 'The Ascent of the Specious', in Cohen 1974.

Bickerton, D. (1972) 'The Structure of Polylectal Grammars', *Georgetown Monograph Series* 25.

(1973a) 'Quantitative versus Dynamic Paradigms: The Case of Montreal *Que*', in Bailey and Shuy 1973.

Bickerton, D. (1973b) 'The Nature of a Creole Continuum', *Lg*, 49.3.

Cedergren, H.J. & Sankoff, D. (1974) 'Variable Rules: Performance as a Statistical Reflection of Competence', *Lg*, 50.2.

Chomsky, N. (1962) 'Explanatory Models in Linguistics', in E.Nagel, P.Suppes & A.Tarski (eds) *Logic, Methodology, and Philosophy of Science*. Stanford U.P.

(1964) *Current Issues in Linguistic Theory*. The Hague, Mouton.

(1965) *Aspects of the Theory of Syntax*. Cambridge (Mass), M.I.T. Press.

Cicourel, A. (1964) *Method and Measurement in Sociology*. Glencoe, Free Press.

(1972) 'Basic and Normative Rules in the Negotiation of Status and Role', in D. Sudnow (ed) *Studies in Social Interaction*. Glencoe, Free Press.

Cohen, D. (ed. 1974) *Explaining Linguistic Phenomena*. Washington, Hemisphere.

Dougherty, R. (1974) 'What Explanation Is and Isn't', in Cohen 1974.

Dretske, F.I. (1974) 'Explanation in Linguistics', in Cohen 1974.

Feys, R. and Fitch, F.B. (1969) *Dictionary of Symbols of Mathematical Logic*. North Holland.

Gazdar, G. (1976a) 'Conversational Analysis and Conventional Sociolinguistics', in J.M.Atkinson and D.R.Watson (eds) *Ethnographics: Studies in Ethnomethodology*. Cambridge, C.U.P.

(1976b) Review of G.Sampson, *The Form of Language*. London, Weidenfeld and Nicolson, 1975. *JL*, 12.1.

Labov, W. (1970) 'The Study of Language in its Social Context', *Studium Generale*, 23.

(1973) 'The Meaning of Words and their Boundaries', in Bailey and Shuy 1973.

Quine, W.V. (1972) 'Methodological Reflections on Current Linguistic Theory', in D.Davidson and G.Harman (eds) *Semantics of Natural Language*. Dordrecht, Reidel.

Sanders, G.A. (1974) Introduction to Cohen 1974.

Sarles, H.B. (1970) 'An Examination of the Question-Response System in Language', *Semiotica*, 2.1.

Schegloff, E. (1968) 'Sequencing in Conversational Openings', *American Anthropologist*, 70.6.

Whitaker, H. (1974) 'Is the Grammar in the Brain?' in Cohen 1974.

Wilson, T.P. (1970) 'Conceptions of Interaction and Forms of Sociological Explanation', *American Sociological Review*, 35.